

Determinants and Dynamics of Migration to OECD Countries in a Three-Dimensional Panel Framework

Ilse Ruyssen, Gerdie Everaert, and Glenn Rayp*

SHERPPA, Ghent University

December 2012

Abstract

This paper investigates the determinants of bilateral immigrant flows to 19 OECD countries between 1998 and 2007 from both advanced and developing origin countries. We pay particular attention to dynamics by including both the lagged migrant flow and the migrant stock to capture partial adjustment and network effects. To correct for the dynamic panel data bias of the fixed effects estimator we use a bootstrap algorithm. Our results indicate that immigrants are primarily attracted by better income opportunities and higher growth rates abroad. Also short-run increases in the host country's employment rate positively affect migration from both advanced and developing countries. High public services, on the other hand, discourage migration from advanced countries but exert a pull on migration from developing sources, in line with the welfare state hypothesis. Finally, we find evidence for both partial adjustment and the presence of strong network effects. This confirms that both should be considered crucial elements of the migration model and that a correction for their joint inclusion is required.

JEL Classification: F22, J61, C33

Keywords: International migration, Network effects, Dynamic panel data model, Bias correction

1 Introduction

Recent changes in both the size and the composition of migrant flows to OECD countries have placed international migration high on the policy agenda in many countries. In terms of size, the number of immigrants residing in the 33 current OECD member states roughly increased from about 42 million in 1980 to over 87 million in 2000. In terms of composition, the expansion of immigration from Central and Eastern Europe to Western Europe following the enlargement of the European Union is apparent, but also migration from India and China to non-European countries has been growing at a steady pace.

*We acknowledge financial support from the Interuniversity Attraction Poles Program - Belgian Science Policy, contract no. P5/21. Corresponding author: Ilse Ruyssen, Department of Economics, Ghent University, Tweakerkenstraat 2, B-9000 Gent, Belgium. Tel.: +32 9 264 42 42. Fax: +32 9 264 35 99. Email: Ilse.Ruyssen@UGent.be.

Understanding the forces that drive such migration patterns is important for the conduct of migration policy.

A general theoretical view on the determinants of migration is the traditional push-pull model (see e.g. Lee, 1966; Todaro, 1969; Borjas, 1989) in which costs and benefits of migrating are determined by push factors of conditions at the origin and pull factors of prospects at the destination. Migration occurs when the net present expected value of migrating is positive. Typical factors are wages and (un)employment rates in both the origin and the destination country, which together determine the expected wage differential. Other factors are levels of social expenditures (Borjas, 1987, 1999; Pedersen et al., 2008; Warin and Svaton, 2008), geographical and cultural proximity (Karemera et al., 2000; Brücker and Siliverstovs, 2006; Lewer and Van den Berg, 2008; Pedersen et al., 2008; Mayda, 2010; Warin and Svaton, 2008) but also differences in living standards and the sociopolitical environment (Karemera et al., 2000; Vogler and Rotte, 2000; Bertocchi and Strozzi, 2008; Hooghe et al., 2008). A popular dynamic factor is given by network effects, which suggests that having friends and family from the same origin living in the host country lowers the monetary and psychological costs of migrating and thus increases migration to that country. As such, migration may become a self-perpetuating process. Surveying empirical findings (see Gould, 1979; Bauer and Zimmermann, 1999, for excellent surveys of the earlier results), the main reason for migration appears to be the search for better economic conditions. Nearly all studies find a significant effect of income differentials between the origin and destination country. The findings regarding (un)employment rates in both sending and receiving countries are more ambiguous, though. Network effects, proxied by including either the lagged migrant flow or the stock of migrants in the destination country, are also found to be very important. The coefficients on these dynamic factors are typically positive and statistically highly significant. Moreover, when excluding such dynamic factors, regression errors are often found to exhibit severe serial correlation. However, there is a longstanding discussion, dating back to e.g. Laber (1972) and Dunlevy and Gemery (1977), on whether these findings signal strong network effects or rather a partial adjustment mechanism reflecting sluggishness in the response of migration to shifts in its underlying determinants. Building on microeconomic utility maximization, Hatton (1995) derives a formal dynamic model of migration in which both the lagged migration flow and the stock of migrants enter as separate determinants, with the former capturing dynamics resulting from uncertainty about future relative income streams and the latter capturing network effects.

Although the empirical literature on migration determinants has made tremendous progress in recent years, it is still plagued by a number of flaws. First, due to data limitations, most studies have estimated the determinants of international migration to a single destination country, either ignoring the origin of migrants, i.e. pure time series models with time as the only dimension (see e.g. Hatton, 1995) or accounting for the origin of migrants using a two-dimensional panel data model with bilateral effects (see Karemera et al., 2000 for immigration to Canada and the US or Vogler and Rotte, 2000; Fertig, 2001; Boeri et al., 2002; Brücker and Siliverstovs, 2006 for the German case). The recent availability of comprehensive data on bilateral migration offers a three-way panel dataset which allows for the inclusion of time dummies next to bilateral effects. This has the important advantage that it allows to control for observed and unobserved time invariant bilateral effects like geographical, historical, political and cultural

influences as well as for time effects like cyclical influences, policy changes, decreases in transportation and communication costs, ..., which are common for all country pairs. Recent studies that estimate a three-way panel data model are Lewer and Van den Berg (2008), Pedersen et al. (2008), Mayda (2010) for immigration to OECD countries or Gallardo-Sejas et al. (2006), Hooghe et al. (2008), Warin and Svaton (2008) for immigration to Europe. Second, none of the above mentioned studies (except Fertig, 2001) allow for the rich migration dynamics present in Hatton's (1995) model, i.e. some studies use a purely static empirical specification while dynamic specifications include either the lagged migrant flow or the stock of migrants but never both of them together, which is required to capture both partial adjustment and network effects. Third, panel datasets on bilateral migration flows and stocks typically hold a small number of time series observations (T) on a moderate number of cross-sections (N). Estimating a dynamic model using such data is particularly challenging. The standard fixed effects (FE) estimator, used by e.g. Hooghe et al. (2008) and Warin and Svaton (2008), is severely biased and inconsistent for T fixed and N going to infinity (see e.g. Nickell, 1981). First-differenced and even system generalized method of moments (GMM) estimators, used by Mayda (2010), are known to suffer from a weak instruments problem (see e.g. Bun and Windmeijer, 2010), which implies a small sample bias, large uncertainty around coefficient estimates and strong sensitivity to instruments choice. We refer to Baltagi (2008) for an overview of dynamic panel data estimators.

In this paper, we investigate the determinants of bilateral immigration to the OECD from both advanced and developing origin countries between 1998 and 2007 using the OECD's International Migration Database. Our contribution is twofold. First, we estimate a dynamic model of migration using a three-way panel data model. In contrast to the literature and in line with Hatton (1995) we include both lagged migration and migration stock which allows us to separately identify network effects and dynamics stemming from partial adjustment. Second, we estimate this dynamic panel data model using an extended version of the bias-corrected fixed effects (BCFE) estimator suggested by Everaert and Pozzi (2007). This estimator corrects for the dynamic panel data bias of the FE estimator using an iterative bootstrap algorithm. Its main advantage over GMM estimators for dynamic panel data is that it combines a small bias with a relatively small standard error. We slightly adjust the bootstrap algorithm of the BCFE estimator to take into account that in our model the dynamic panel data bias is induced by the lagged migrant flow as well as by the migrant stock. Using Monte Carlo experiments, we demonstrate that this adjusted BCFE estimator performs well in the specific context of our model and is preferable to alternative estimators.

Our results indicate that immigrants are primarily attracted by better income opportunities abroad and much less by income at home and by employment rates both at home and abroad. High public services are found to discourage migration from advanced countries but exert a pull on migration from developing sources, confirming the welfare magnet hypothesis. Furthermore, we find evidence of strong dynamic effects. Both the lagged migration flow and the migrant stock have a strong positive and significant impact on current migration, the former indicating dynamic effects stemming from the process by which expectations about future earnings are formed and updated while the latter indicates network effects. Further evidence that dynamics play a prominent role in the migration model arises from the observation

that misspecifying the model by omitting the lagged migration flow or the migrant stock and/or not correcting for the dynamic panel bias has a strong impact on the estimation results. Therefore, care should be taken when specifying the dynamic structure of the model and selecting the estimation method.

The remainder of this paper is organized as follows. Section 2 derives the empirical specification and presents the estimation method with Monte Carlo evidence on its performance. Section 3 describes the data and reports the estimation results. Section 4 summarizes the major findings.

2 A three-way dynamic panel data approach to migration

One of the major contributions to the literature on the determinants of migration has been the traditional push-pull model (see e.g. Lee, 1966; Todaro, 1969; Borjas, 1989). According to this model, migration is the result of push factors at the origin and pull factors at the destination. The migration decision is based on the comparison between expected benefits and costs of migration. A formal dynamic model was developed by Hatton (1995). This model forms the basis for our empirical specification.

2.1 A dynamic model of migration

Hatton's model builds on a microeconomic analysis which treats migration as a decision of a utility maximizing individual. The probability of migration depends on the difference in expected utility streams between the origin (o) and the destination (d) country. For an individual i , this difference in year t is given by

$$d_{it} = Eu(y_{dt}) - Eu(y_{ot}) + z_{it}, \quad (1)$$

where y is income and z_i captures the individual's non-pecuniary utility difference between the two countries and the cost of migration. Following Todaro (1969), Hatton defines expected income as the wage (w) times the employment rate (e), with income uncertainty being due to uncertain employment prospects. To take into account the welfare magnet theory presented in Borjas (1987, 1999), we extend this definition of expected income by adding the provision of public services (ps) in the form of social protection benefits¹ (see also Pedersen et al., 2008; Warin and Svaton, 2008). Assuming a logarithmic utility function and a binomial distribution to characterize the probability of employment, equation (1) can be rewritten as

$$d_{it} = \eta_1 \ln w_{dt} - \eta_2 \ln w_{ot} + \eta_3 \ln ps_{dt} + \eta_4 \ln e_{dt} - \eta_5 \ln e_{ot} + z_{it} \quad (2)$$

A key dynamic feature of migration is that the decision to migrate does not only depend on the current utility difference, d_{it} , but also on the net present value of all future utility differences, denoted d_{it}^* . Hence, the total net present value of migrating today is given by $d_{it} + d_{it}^*$. Moreover, even if today's total net present value is positive, it might even be higher in the future (i.e. $d_{it}^* > d_{it} + d_{it}^*$) when $d_{it} < 0$,

¹The inclusion of public services might also be linked to the cost of migration, z_{it} . In that sense, immigrants are expected to prefer countries with a generous system of public services since the presence of a safety net lowers the psychological cost of migration.

which makes it interesting for the individual to wait. Consequently, the probability of migrating at time t ($m_{it} = 1$) is given by

$$Pr(m_{it} = 1) = Pr(d_{it} + d_{it}^* > 0 \cap d_{it} > 0). \quad (3)$$

To capture this, Hatton writes aggregate migration as

$$\ln M_{dot} = \beta(d_t^* + \alpha d_t) = \beta d_t^* + \beta \alpha d_t \quad (4)$$

where M_{dot} is the aggregate migration flow from origin country o to destination country d at time t and $\alpha > 1$ reflects the extra weight given to current conditions. Assuming that expectations about future utility streams are a geometric series of past utility differences²

$$d_t^* = \lambda d_t + \lambda^2 d_{t-1} + \lambda^3 d_{t-2} + \lambda^4 d_{t-3} + \dots, \quad (5)$$

and using the Koyck transformation, this results in the following aggregate dynamic migration equation

$$\begin{aligned} \ln M_{dot} = & \lambda \ln M_{dot-1} + \beta(\alpha + \lambda) [\eta_1 \ln w_{dt} - \eta_2 \ln w_{ot} + \eta_3 \ln ps_{dt} + \eta_4 \ln e_{dt} - \eta_5 \ln e_{ot} + \bar{z}_t] \\ & - \lambda \beta \alpha [\eta_1 \ln w_{dt-1} - \eta_2 \ln w_{ot-1} + \eta_3 \ln ps_{dt-1} + \eta_4 \ln e_{dt-1} - \eta_5 \ln e_{ot-1} + \bar{z}_{t-1}]. \end{aligned} \quad (6)$$

Hatton assumes that \bar{z} , which is the mean of z_i , is determined by the stock of previous immigrants and a time variable such that

$$\bar{z}_t = \gamma_0 + \gamma_1 \ln MST_{dot} + \gamma_t + \gamma_{do} \quad (7)$$

where MST_{dot} is the stock of migrants from origin country o residing in destination country d at the beginning of time t . This stock variable is included to capture network effects: friends and relatives who already live in the host country reduce the monetary and psychological costs of migration. The higher the stock of previous immigrants from the same origin country, the lower the costs of migration and the higher the immigrant flow. Nevertheless, this is not the only cost determining factor. Also decreasing transportation and communication costs lower the cost of migration over time. In our model, these decreasing costs are captured by the year dummies γ_t . The latter might however also represent, among other things, the impact of joint changes in origin and destination countries' emigration and immigration policies. Furthermore, also distance, common language, similar culture, colonial ties and immigration policy affect the cost of migration. To the extent that these factors are time invariant, they are captured by the bilateral fixed effect γ_{do} .

The stock of migrants diminishes at a rate δ_{do} due to deaths and return migration but increases due to the inflow of new migrants such that

$$MST_{dot} = (1 - \delta_{do}) MST_{dot-1} + M_{dot-1}. \quad (8)$$

where δ_{do} is allowed to vary across destination and origin country pairs. In a later stage, this relationship will be used to account for the link between immigrant flows and stocks. For the moment, we use this

²As shown by Hatton (1995) this is consistent with rational expectations if d_{it} follows an AR(1) process.

expression to eliminate $\ln MST_{dot-1}$ from \bar{z}_{t-1} in equation (6) by applying a logarithmic expansion of the migrant stock and its components in equation (8) about their mean values so that

$$\ln MST_{dot} = (1 - \Omega) \ln [(1 - \delta_{do}) MST_{dot-1}] + \Omega \ln M_{dot-1}, \quad (9)$$

where $\Omega = \frac{M}{(1-\delta)MST+M} > 0$.³

Substituting (7) and (9) in (6) and rearranging gives

$$\begin{aligned} \ln M_{dot} = & \mu_0 + \mu_t + \mu_{do} + \left(1 - \frac{\Omega\beta\alpha\gamma_1}{1-\Omega}\right) \lambda \ln M_{dot-1} + \left(\beta(\alpha + \lambda) - \frac{\beta\alpha\lambda}{1-\Omega}\right) \gamma_1 \ln MST_{dot} \\ & + \beta(\alpha + \lambda) \eta_1 \ln w_{dt-1} - \beta(\alpha + \lambda) \eta_2 \ln w_{ot-1} + \beta(\alpha + \lambda) \eta_3 \ln ps_{dt-1} \\ & + \beta(\alpha + \lambda) \eta_4 \ln e_{dt-1} - \beta(\alpha + \lambda) \eta_5 \ln e_{ot-1} \\ & + \beta(\alpha + \lambda - \alpha\lambda) \eta_1 \Delta \ln w_{dt} - \beta(\alpha + \lambda - \alpha\lambda) \eta_2 \Delta \ln w_{ot} + \beta(\alpha + \lambda - \alpha\lambda) \eta_3 \Delta \ln ps_{dt} \\ & + \beta(\alpha + \lambda - \alpha\lambda) \eta_4 \Delta \ln e_{dt} - \beta(\alpha + \lambda - \alpha\lambda) \eta_5 \Delta \ln e_{ot} + \varepsilon_{dot} \end{aligned} \quad (10)$$

with $\mu_0 = \beta(\alpha + \lambda - \alpha\lambda) \gamma_0$, $\mu_t = \beta(\alpha + \lambda) \gamma_t - \beta\alpha\lambda\gamma_{t-1}$ and $\mu_{do} = \beta(\alpha + \lambda - \alpha\lambda) \gamma_{do} + \beta\lambda\alpha\gamma_1 \ln(1 - \delta_{do})$.

A number of key features of this model are worth discussing. First, note that equation (10) is of the double log form, which results from the choice of functional form for the utility function and from taking migration and the migrant stock in equations (4) and (7) as logarithmic. Although Hatton's (1995) original model is semi-logarithmic, he emphasizes that this model is only one among many different functional forms and also suggests and estimates a double log version. Given our panel dataset, with countries that greatly differ in size, the double log form has the important advantage that it eliminates the scale of the migrant flows and stocks. As an alternative, some studies divide the immigrant flow by the population in the origin or destination countries (see e.g. Fertig, 2001; Boeri et al., 2002; Pedersen et al., 2008; Mayda, 2010), but this only partly removes problems of scale. Only dividing by the population in both sending and receiving countries or taking the natural logarithm entirely solves the problem (see Lewer and Van den Berg, 2008; Warin and Svaton, 2008; Ortega and Peri, 2009).

Second, lagged migration flow and migrant stock enter equation (10) as two separate determinants. This contradicts the common practice in empirical studies to include either the lagged migration flow (see e.g. Bertocchi and Strozzi, 2008; Mayda, 2010) or the migrant stock (see e.g. Hooghe et al., 2008; Lewer and Van den Berg, 2008; Pedersen et al., 2008; Warin and Svaton, 2008), with both variables typically being argued to capture network effects⁴. Laber (1972) has already highlighted that it is not clear whether these dynamic terms represent network effects or rather capture a partial adjustment

³First, we can write $\ln \{MST_{dot} / [(1 - \delta_{do}) MST_{dot-1}]\}$ as $\ln \{1 + \exp[\ln M_{dot-1} - (1 - \delta_{do}) \ln MST_{dot-1}]\}$. A first-order Taylor expansion of the latter around the mean values of M_{dot-1} and $(1 - \delta_{do}) MST_{dot-1}$ gives $\ln \{MST_{dot} / [(1 - \delta_{do}) MST_{dot-1}]\} \approx \Omega [\ln M_{dot-1} - (1 - \delta_{do}) \ln MST_{dot-1}] + c$ where c is an arbitrary constant which we ignore for notational convenience. Now add $\ln [(1 - \delta_{do}) MST_{dot-1}]$ to both sides of the equation to approximate $\ln MST_{dot} = \ln [(1 - \delta_{do}) MST_{dot-1} + M_{dot-1}]$ which gives (9) in the text.

⁴A popular motivation for not including both lagged migrant flow and migrant stock is that the latter is, as presented in equation (8), the sum of all past immigrant flows less deaths and return migrants. Hence, the migrant stock is itself a function of all those factors which influenced the earlier immigrant flows. Therefore it will be correlated with all the explanatory variables. However, multicollinearity is no reason to omit the migrant stock variable as this may result in a specification bias as well as in a loss of information regarding the network effect.

mechanism. Dunlevy and Gemery (1977) argue that lagged migration and migrant stock should both be included as determinants to capture the separate impact of partial adjustment and network effects respectively. This is confirmed by equation (10) which shows that a nonzero coefficient on $\ln M_{dot-1}$ implies partial adjustment ($\lambda \neq 0$) stemming from the process by which expectations are formed and updated while a nonzero coefficient on $\ln MST_{dot}$ implies network effects ($\gamma_1 \neq 0$).

Third, an additional dynamic feature of the model is that it includes both lagged levels and current changes of the explanatory variables. The latter capture immediate responses of the immigrant flow to changes in the explanatory variables. This stems from the fact that migration decisions can be postponed when economic conditions are unfavorable such that migration may fluctuate more closely with current conditions than might be expected from individuals that maximize their lifetime utilities.

2.2 Empirical specification and long-run effects

The unrestricted form of equation (10) is given by

$$\begin{aligned}\ln M_{dot} &= \mu_{do} + \mu_t + \theta_1 \ln M_{dot-1} + \theta_2 \ln MST_{dot} + \theta_3 X_{dot-1} + \theta_4 \Delta X_{dot} + \varepsilon_{dot}, \\ &= \mu_{do} + \mu_t + \theta W'_{dot} + \varepsilon_{dot},\end{aligned}\tag{11}$$

where $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ and $W_{dot} = (\ln M_{dot-1}, \ln MST_{dot}, X_{dot-1}, \Delta X_{dot})$ with X_{dot} capturing all determinants of migration other than $\ln M_{dot-1}$ and $\ln MST_{dot}$ included in equation (10). The error terms ε_{dot} are assumed to be serially uncorrelated but allowed to be heteroscedastic between and contemporaneously correlated over cross-sections. Both the lagged migration flow $\ln M_{dot-1}$ and the migrant stock $\ln MST_{dot}$ (which is measured at the beginning of time t) are predetermined at time t and therefore not correlated with the error term ε_{dot} . By construction, both these variables are correlated with the individual effects μ_{do} . All other regressors are allowed to be correlated with μ_{do} but are assumed to be exogenous with respect to ε_{dot} . The latter assumption is based on the fact that we investigate the determinants of bilateral immigrant flows, i.e. at a disaggregated level, which will have only a small impact on the macroeconomic determinants of migration like e.g. wages and employment. The semi long-run impact of the explanatory variables on migrant flows can be obtained by imposing a no change constraint on equation (11), i.e. imposing $\ln M_{dot} = \ln M_{dot-1}$, $\varepsilon_{dot} = 0$ and setting differences to zero, which gives

$$\ln M_{dot} = \frac{1}{1 - \theta_1} (\mu_{do} + \mu_t + \theta_2 \ln MST_{dot} + \theta_3 X_{dot}).\tag{12}$$

Yet, these are not the full long-run effects as they ignore the endogeneity of the migrant stock $\ln MST_{dot}$. The full long-run impact is obtained by simulating the dynamic response⁵ of $\ln M_{do}$ and $\ln MST_{do}$ to a 1% increase in each of the explanatory variables in X_{do} using the estimated equation (11) together with equation (8) with $\delta = \frac{1}{N} \sum_{do} \delta_{do}$.

⁵The long-run impact is defined from imposing a no change condition, i.e. the criterion that the squared difference between two subsequent values of the dynamic response should be less than or equal to 0.0001².

2.3 Choice of dynamic panel data estimator

The empirical specification in equation (11) is dynamic in the sense that it incorporates both the lagged migrant flow and the migrant stock as explanatory variables. Estimation of dynamic panel data models has received a lot of attention in the literature. The main problem is that the lagged dependent variable is by construction correlated with the individual effects. This renders the pooled ordinary least-squares (POLS) estimator biased and inconsistent. A within transformation wipes out the individual effects by taking deviations from individual sample means, but the resulting FE estimator is biased and inconsistent for fixed T and N going to infinity (see Nickell, 1981). Given this inconsistency, the literature focuses mainly on a first-difference transformation to eliminate the individual effects while handling the remaining correlation with the (transformed) error term using instrumental variables (IV) and GMM estimators. Especially the first-differenced GMM estimator of Arellano and Bond (1991) and the system GMM estimator of Arellano and Bover (1995) and Blundell and Bond (1998) are popular. The advantage of these estimators is that they are consistent for fixed T and large N . Unfortunately, these GMM estimators (i) have a (much) larger standard error compared to the FE estimator (see e.g. Arellano and Bond, 1991; Kiviet, 1995) and (ii) may suffer from a substantial finite sample bias due to weak instrument problems (see Ziliak, 1997; Bun and Kiviet, 2006; Bun and Windmeijer, 2010). In order to avoid these problems, analytical bias-corrections for the FE estimator have been proposed by, among others, Kiviet (1995), Bun (2003) and Bun and Carree (2005). The advantage of these estimators is that they reduce the bias of the FE estimator while maintaining its small dispersion relative to GMM. Although these estimators perform remarkably well, even in samples of moderate T , the use of analytical corrections in practical applications may be limited as the theoretical restrictions under which these corrections are derived do not necessarily hold. A detailed overview of dynamic panel data estimators can be found in Baltagi (2008).

As an alternative, Everaert and Pozzi (2007) propose a bias-correction for the FE estimator using an iterative bootstrap algorithm. Like analytical corrections, this bootstrap correction reduces the bias of the FE estimator while maintaining its higher efficiency compared to GMM estimators. The main advantage is that it can more easily be adjusted to practical applications by an appropriate choice of the data resampling scheme. This flexibility is of particular interest for estimating our empirical specification where next to the lagged migration flow also the migration stock is by construction correlated with the individual effects. This is a case which is not considered by the analytical corrections. Therefore, the bootstrap-based bias-corrected FE estimator is our main estimator used below. We refer to it as BCFE.

2.4 Implementation of the BCFE estimator

Without going in too much technical details (for this we refer to Everaert and Pozzi, 2007), the basic BCFE estimator searches over the parameter space and takes as bias-corrected estimates the set of parameters $\tilde{\theta}$ for which holds that when repeatedly generating artificial data from equation (11) setting $\theta = \tilde{\theta}$ and next estimating this equation from these artificial data using FE yields on average (over repeated samples) the original biased FE estimates $\hat{\theta}$. In practice, this search over the parameter space is computationally implemented through an iterative bootstrap algorithm, initiated by setting as a first guess $\tilde{\theta}^0 = \hat{\theta}$, which

is used to generate 1000 bootstrap data samples from equation (11) setting $\theta = \tilde{\theta}^0$. These artificial data are then used to calculate the bias of the FE estimator as $\tilde{\theta}^0 - \bar{\theta}^1$ where $\bar{\theta}^1$ is the average of the 1000 FE estimates obtained over the bootstrap samples. The first step bias-corrected FE estimator is then given by $\tilde{\theta}^1 = \hat{\theta} + (\tilde{\theta}^0 - \bar{\theta}^1)$. In the second step, this bias-correction procedure is repeated but now data are generated by setting $\theta = \tilde{\theta}^1$ from which we obtain the bias as $\tilde{\theta}^1 - \bar{\theta}^2$ (with obvious notation) and the second step bias-corrected FE estimator as $\tilde{\theta}^2 = \hat{\theta} + (\tilde{\theta}^1 - \bar{\theta}^2)$. This procedure is then iterated until convergence, i.e. a stable set of parameter values $\tilde{\theta}^k \approx \tilde{\theta}^{k+1}$ is obtained.

The artificial data generated in the algorithm outlined above are obtained using a semi-parametric procedure, i.e. bootstrap samples $\tilde{\varepsilon}_{dot}^b$ are obtained by a non-parametric resampling of the (rescaled) estimated residuals $\tilde{\varepsilon}_{dot}^k$ (obtained using $\tilde{\theta}^k$) while bootstrap samples for M_{dot}^b are calculated from the parametric model in equation (11) setting $\theta = \tilde{\theta}^k$. As stated above, this data resampling procedure has the important advantage that it can easily be shaped to align with the assumed data generating process of the data. First, the non-parametric resampling of $\tilde{\varepsilon}_{dot}^k$ does not require explicit distributional assumptions for the population errors ε_{dot} such that, in line with our assumptions in section 2.2, we allow for (i) heteroscedasticity over cross-sections by resampling residuals within but not between cross-section units and (ii) contemporaneous correlation between cross-sections by applying the same resampling index to each cross-section. Second, next to calculating bootstrap samples M_{dot}^b for the migrant flow from equation (11), we also calculate bootstrap samples MST_{dot}^b for the migrant stock using equation (8) setting $\delta_{do} = 1 - \frac{1}{T} \sum_t ((MST_{dot} - M_{dot-1}) / MST_{dot-1})$. This captures the important feature that MST_{dot} is endogenous, i.e. correlated with the individual effects. Further note that in line with the assumption in section 2.2 that all explanatory variables other than M_{dot} and MST_{dot} are exogenous, these are kept fixed over the bootstrap samples.

2.5 Monte Carlo simulation

In this section we conduct a small-scaled Monte Carlo experiment to assess the finite sample properties of our adjusted BCFE estimator compared to several other estimators.

Design

The data generating process (DGP) is chosen such that the properties of the simulated data match with those of the observed data as much as possible:

- The sample size of the simulated data equals the one available for estimation. This implies running separate simulations for advanced ($T = 9, N = 247$) and developing ($T = 9, N = 388$) origin countries.
- Data for the endogenous variables migration flow M_{dot} and migration stock MST_{dot} are drawn from their data generating process (DGP) in equations (11) and (8) respectively, using the observed values in the first year of the sample as initialisation.
- The parameter values for θ in the DGP for M_{dot} in equation (11) are set equal to the BCFE

estimates from Table 3 below while δ_{do} in equation (8) is set equal to the value observed in the sample data.

- Error terms ε_{dot} are generated from a normal distribution with estimated variance from the residuals of the BCFE regressions in Table 3.
- The observed values for the exogenous variables X_{dot-1} and ΔX_{dot} are treated as fixed in each MC iteration.

We generate data both for the full model and for a partial model with only stocks and lagged flows as explanatory variables ($\theta_3 = \theta_4 = 0$ in equation (11)). This results in four experiments with the coefficients for lagged flows, θ_1 , and stocks, θ_2 , respectively set to 0.61 and 0.46 (0.64 and 0.49) for the complete (partial) model using the advanced dataset, and 0.75 and 0.23 (0.74 and 0.23) for the complete (partial) model using the developing dataset. In each experiment, we perform 1000 replications.

Estimators

We compare the performance of the BCFE estimator with (i) FE, the standard fixed effects estimator, (ii) GMMd, the first-difference GMM estimator proposed by Arellano and Bond (1991) and (iii) GMMs, the system GMM estimator proposed by Arellano and Bover (1995) and Blundell and Bond (1998). For the GMMd estimator, at least one period lagged values ($\ln M_{dot-1-s}$ and $\ln MST_{dot-s}$ with $s \geq 1$) are available as instruments for the predetermined variables $\ln M_{dot-1}$ and $\ln MST_{dot}$ ⁶ in each period. For the exogenous variables X_{dot-1} and ΔX_{dot} , the available instruments set is $(X_{do1}, \dots, X_{doT-1}, \Delta X_{do2}, \dots, \Delta X_{doT})$ in each period. GMMs has the same instrument set as GMMd in the first difference part of the system and has $\Delta \ln M_{dot-2}$, ΔX_{dot-1} and $\Delta^2 X_{dot}$ as additional instruments in the levels part of the system. Note that the first-differenced stock $\Delta \ln MST_{dot-1}$ can not be used as instrument as it is by construction correlated with the fixed effect μ_{do} in the levels equation. Given the large number of exogenous variables, we try to avoid an overfitting bias resulting from using too many instruments (see Ziliak, 1997; Arellano, 2003) by (i) only using the first three available instruments for the predetermined variables ($\ln M_{dot-1}$ and $\ln MST_{dot}$) and the contemporaneous values for the exogenous variables and (ii) stacking the instrument matrix as suggested by Roodman (2009). We report both one-step and two-step GMM estimates.

Simulation results

The simulation results are presented in Tables 1 and 2. For each estimator, we report mean bias, standard deviation (std) and root mean squared error (rmse) in estimating θ_1 and θ_2 .

First looking at the performance in estimating θ_1 , we observe the following results for both types of models. As expected, the FE estimator is biased downward because of the correlation between the transformed lagged dependent variable and the transformed error term. Correcting for the dynamic panel bias by performing BCFE significantly reduces the bias while maintaining the low dispersion associated with the uncorrected FE. The bias of the GMMd1 and GMMd2 estimators is of the same order as in BCFE, but they have a much larger dispersion and rmse. The GMMs estimators have a sizable bias in

⁶ $\ln MST_{dot}$ is predetermined as it is defined as the migrant stock at the beginning of the period.

all cases. This suggests that the extra moment conditions imposed in the level part of the system, from a restriction on the initial conditions process generating $\ln M_{dot1}$, is violated.

Table 1: Monte Carlo results based on database with advanced origins ($T = 9, N = 247$)

	Bias θ_1	Std θ_1	Rmse θ_1	Bias θ_2	Std θ_2	Rmse θ_2
<i>Full model, $\theta_1 = 0.61$ and $\theta_2 = 0.46$</i>						
FE	-0.192	0.021	0.193	0.035	0.047	0.059
BCFE	-0.011	0.025	0.028	-0.026	0.045	0.052
GMMd1	-0.014	0.138	0.138	0.003	0.064	0.064
GMMd2	-0.014	0.141	0.141	0.002	0.065	0.065
GMMs1	0.226	0.026	0.228	-0.249	0.031	0.251
GMMs2	0.200	0.031	0.202	-0.204	0.042	0.208
<i>Partial model with $\theta_3 = \theta_4 = 0$, $\theta_1 = 0.64$ and $\theta_2 = 0.49$</i>						
FE	-0.185	0.021	0.186	0.079	0.044	0.091
BCFE	-0.011	0.025	0.028	-0.024	0.043	0.049
GMMd1	-0.006	0.076	0.076	0.002	0.066	0.066
GMMd2	-0.005	0.076	0.076	0.001	0.066	0.066
GMMs1	-0.112	0.070	0.132	-0.090	0.062	0.109
GMMs2	-0.099	0.071	0.122	-0.103	0.065	0.122

Notes: θ_1 and θ_2 denote the coefficients for $\ln M_{dot-1}$ and $\ln MST_{dot}$, respectively. θ_3 and θ_4 represent the coefficients of the strictly exogenous variables. For the GMM estimators, ‘1’ refers to one-step estimates and ‘2’ refers to two-step estimates.

Second, regarding the relative performance in the estimation of θ_2 , the GMMd estimators have the smallest bias, followed by the FE and BCFE estimators. However, the standard deviation of the GMMd estimators is always bigger compared to the FE and BCFE estimators. This results in (i) the lowest rmse for the BCFE estimator using the advanced dataset and the partial model developing dataset and (ii) a fairly similar rmse for the BCFE and GMMd estimates for the complete model developing dataset. The GMMs estimators again have a sizable bias in most cases.

In conclusion, due to its small bias combined with a relatively small standard deviation, the BCFE estimator is shown to outperform the alternative estimators in terms of rmse given the specificities of our model and sample data. As such, we take it as our preferred estimator in the next section.

3 Data and estimation results

3.1 Data

Data on bilateral immigrant flows and stocks are taken from the International Migration Database provided by the OECD. It contains information on inflows of foreigners by nationality and stocks of foreigners by both nationality and country of birth to 19 OECD countries from 189 origin countries over the period

Table 2: Monte Carlo results based on database with developing origins ($T = 9, N = 388$)

	Bias θ_1	Std θ_1	Rmse θ_1	Bias θ_2	Std θ_2	Rmse θ_2
<i>Full model, $\theta_1 = 0.75$ and $\theta_2 = 0.23$</i>						
FE	-0.159	0.016	0.159	-0.117	0.048	0.127
BCFE	-0.018	0.018	0.025	-0.099	0.044	0.109
GMMd1	-0.024	0.152	0.154	-0.016	0.104	0.105
GMMd2	-0.022	0.153	0.155	-0.017	0.106	0.107
GMMs1	0.162	0.020	0.163	-0.212	0.027	0.214
GMMs2	0.137	0.023	0.139	-0.181	0.032	0.184
<i>Partial model with $\theta_3 = \theta_4 = 0$, $\theta_1 = 0.74$ and $\theta_2 = 0.23$</i>						
FE	-0.136	0.015	0.137	0.039	0.037	0.054
BCFE	-0.007	0.017	0.018	-0.028	0.035	0.045
GMMd1	-0.000	0.042	0.042	-0.002	0.055	0.055
GMMd2	-0.000	0.042	0.042	-0.002	0.056	0.056
GMMs1	0.001	0.042	0.042	-0.013	0.034	0.036
GMMs2	0.000	0.042	0.042	-0.013	0.034	0.036

Notes: see Table 1.

1998-2007. For the migrant stock, we use data on foreign-born by country of birth wherever possible and foreign nationals otherwise. In order to account for potential heterogeneity, we divide our sample of origins into advanced and developing countries following IMF definitions. While the IMF distinguishes between advanced countries on the one hand and developing and emerging countries on the other hand, we combine the second group and refer to it as developing countries. Table A-1 reports total yearly immigrant flows into each destination country between 1998 and 2007. After removing cross-sections with missing observations and with obvious inconsistencies between flows and stocks, we have 247 cross-sections for advanced origins and 388 cross-sections for developing origins. Tables A-2 and A-3 show that these account for 16.5 percent and 58.5 percent respectively of the total flow. Hence, migration from developing countries clearly dominated during our sample period.

Table A-4 presents descriptive statistics for all variables used in the regression analysis. Due to a lack of real wage data for the set of origin countries, wages are approximated by per capita gross domestic product (see also Fertig, 2001; Pedersen et al., 2008; Mayda, 2010), expressed in current dollars purchasing power parities to correct for differences in the evolution in the cost of living between countries. Data on GDP per capita are taken from the Penn World Tables 6.3. The employment rate is proxied by the number of employed relative to the population, as provided by the United Nations Statistics Division. One could argue that the general employment rate does not capture the true labor market constraints faced by immigrants due to the presence of a home bias in the demand for labor. One possibility is to replace it by the employment rate for foreigners in the destination country. However, this rate does not eliminate measurement error since it does not discriminate between foreigners from the developing world

and those from advanced countries. Consequently, we stick to the general employment rate to proxy for employment possibilities for immigrants in the host country. Public services in the destination country are proxied by expenditures of social protection benefits for sickness/health care and family/children allowances, expressed as a percentage of GDP. Generally public expenditures include also other types of benefits such as those for disability, old age, unemployment or housing. Yet, access to those benefits for new entrants is typically constrained and therefore excluded from our proxy for public expenditures. The data on public expenditures were obtained from the Social Expenditure Database (SOCX), provided by the OECD.

3.2 Estimation results

The estimation results are reported in Table 3. To allow for a heterogeneous impact of migration determinants, separate results are reported for migration from advanced and from developing countries. Our preferred methodology is BCFE estimation of equation (11). The standard errors used to calculate the t -statistics are simulated using the bootstrap algorithm as outlined in Everaert and Pozzi (2007). They are robust to both cross-sectional heteroscedasticity and cross-sectional error correlation. To link our results to those in the literature, we also report results from (i) FE estimation of restricted versions of equation (11) including either lagged migration or the migrant stock and (ii) FE estimation of equation (11) not correcting for the dynamic panel data bias. For these estimators, standard errors are simulated in a similar way as for the BCFE estimator.⁷

We also experimented with GMMd and GMMs estimations but these were unsatisfactory as the results were highly sensitive to the choice of instruments. Consequently, we do not discuss the GMM results but some of the results can be found in Table A-5 in the Appendix. One interesting point to note though is that, in line with the results from the Monte Carlo simulation, the Sargan-Hansen test rejects the validity of the moment conditions underlying the GMMs estimator. Furthermore, we tested if the model specification in equation (11) is appropriate by adding the second lag of $\ln M_{dot}$ to the estimation equation. The coefficient for the second lag of $\ln M_{dot}$ turned out insignificant for both advanced and developing origins, yet the first lag remained significant indicating that our results are robust for this alternative specification⁸.

Table 4 reports long-run elasticities of migration determinants calculated from the BCFE estimation results. The first three columns report semi long-run effects, while the last three columns report full long-run effects. With respect to the latter, it should be noted that they are calculated assuming the strong link between flows and stocks as given in equation (8). In our dataset this link is less strong, though, as stock data are not constructed from the flow data such that the evolution in flows and stocks is not fully compatible. The exact numbers of the full long-run effects reported in Table 2 should therefore be interpreted with care. Standard errors for the long-run effects are also simulated using the bootstrap algorithm. In line with Everaert and Pozzi (2007), we report the median and the 5th and 95th percentiles of the simulated distribution of the long-run effects rather than the mean and the t -statistic. The reason

⁷The matlab code for the BCFE estimator is available upon request.

⁸The estimation results for this model are available upon request.

Table 3: Estimation results

Dependent variable: $\ln M_{dot}$					Sample period: 1998-2007			
	Advanced origins				Developing origins			
	FE(1)	FE(2)	FE(3)	BCFE	FE(1)	FE(2)	FE(3)	BCFE
$\ln MST_{dot}$		0.73*** (6.19)	0.44*** (4.82)	0.46*** (6.10)		0.82*** (7.75)	0.24*** (4.54)	0.23*** (4.55)
$\ln M_{dot-1}$	0.48*** (9.38)		0.44*** (7.58)	0.61*** (8.51)	0.65*** (27.67)		0.61*** (24.02)	0.75*** (13.50)
$\ln w_{dt-1}$	0.98*** (3.45)	0.93** (2.00)	0.70*** (2.59)	0.59** (2.32)	1.92*** (5.82)	2.78*** (4.08)	1.82*** (5.43)	1.58*** (4.89)
$\ln w_{ot-1}$	-0.30 (-1.53)	-0.29 (-1.00)	-0.34* (-1.75)	-0.37** (-2.20)	0.08 (0.81)	0.15 (0.82)	0.05 (0.53)	0.00 (0.02)
$\ln ps_{dt-1}$	-0.24 (-1.39)	-0.39 (-1.58)	-0.41** (-2.40)	-0.49*** (-2.95)	0.66** (2.22)	0.86* (1.68)	0.58* (1.86)	0.37 (1.19)
$\ln e_{dt-1}$	1.06* (1.81)	2.56*** (3.11)	0.96 (1.63)	0.26 (0.46)	-0.29 (-0.58)	1.86** (2.12)	-0.69 (-1.45)	-1.56** (-2.30)
$\ln e_{ot-1}$	0.10 (0.27)	0.31 (0.54)	0.30 (0.79)	0.29 (0.84)	-0.18 (-0.75)	-0.30 (-0.69)	-0.17 (-0.72)	-0.17 (-0.78)
$\ln \Delta w_{dt}$	1.34*** (2.83)	1.56** (2.37)	1.40*** (2.75)	1.51*** (2.85)	2.07*** (4.35)	2.47*** (3.22)	2.21*** (4.82)	2.45*** (5.00)
$\ln \Delta w_{ot}$	-0.02 (-0.08)	0.16 (0.46)	-0.06 (-0.24)	-0.22 (-0.85)	0.12 (0.99)	0.09 (0.56)	0.10 (0.82)	0.08 (0.59)
$\ln \Delta ps_{dt}$	0.52** (2.16)	0.60** (1.99)	0.55** (2.32)	0.55** (2.06)	1.08*** (3.30)	0.82* (1.64)	1.15*** (3.72)	1.33*** (3.96)
$\ln \Delta e_{dt}$	2.31*** (3.35)	3.14*** (3.79)	2.53*** (3.84)	2.18*** (3.12)	2.82*** (3.13)	1.90* (1.82)	2.35*** (2.62)	1.89* (1.79)
$\ln \Delta e_{ot}$	1.26* (1.87)	2.11*** (2.55)	1.10* (1.65)	0.76 (1.13)	0.13 (0.37)	0.34 (0.69)	0.14 (0.40)	0.05 (0.15)

Notes: Each regression includes time dummies (not reported). t -statistics - between brackets - are robust to cross-sectional heteroskedasticity and cross-sectional error correlation. *, ** and *** indicate significance at the 10%, 5% and 1% level respectively. *Advanced*: 2223 observations and 247 cross sections. *Developing*: 3492 observations and 388 cross sections.

for this is that the distribution of the long-run effects does not necessarily have finite moments, especially when the root of the dynamic process is close to unity. It should be noted that these percentiles are not necessarily finite either but they should be less vulnerable to large outliers in the distribution.

3.2.1 Dynamic features of migration

Consistent with the findings in for instance Fertig (2001), Clark et al. (2002) and Pedersen et al. (2008), lagged migrant flows and migrant stocks appear to have the most pervasive impact on subsequent migration from both advanced and developing countries. The results from our preferred BCFE estimator suggest an elasticity of 0.61 (0.75) for lagged migrant flows from advanced (developing) countries and 0.46 (0.23) for the stock of migrants from advanced (developing) countries. The fact that both are significant indicates that multicollinearity between these two variables is fairly small. In correspondence to earlier findings in (Dunlevy and Gemery, 1977) it seems that these variables do not measure the same phenomenon, supporting their simultaneous inclusion in the estimation equation. The significant coefficient on lagged migration flows suggests dynamic effects stemming from the process by which expectations

Table 4: Long-run estimation results

Dependent variable: $\ln M_{do}$				Sample period: 1998-2007		
	Semi LR (BCFE)			Full LR (BCFE)		
	median	percentiles		median	percentiles	
		5th	95th		5th	95th
<i>Advanced origins</i>						
$\ln MST_{dot}$	0.93***	0.68	1.25	0.00	0.00	0.00
$\ln w_{dt}$	1.64**	0.52	3.26	5.32***	1.72	2.27
$\ln w_{ot}$	-0.95*	-1.79	-0.19	-2.00**	-3.26	-3.06
$\ln ps_{dt}$	-1.19***	-2.00	-0.48	-2.31***	-3.60	-3.33
$\ln e_{dt}$	1.19	-1.81	3.03	3.67	-3.15	-2.15
$\ln e_{ot}$	1.02	-0.66	2.40	2.23	-1.46	-0.69
<i>Developing origins</i>						
$\ln MST_{dot}$	1.01***	0.56	1.54	0.00	0.00	0.00
$\ln w_{dt}$	7.00***	4.36	10.31	14.82***	7.62	28.57
$\ln w_{ot}$	0.04	-0.76	0.56	0.15	-1.12	0.79
$\ln ps_{dt}$	1.93	0.02	4.65	3.17*	0.29	8.44
$\ln e_{dt}$	-7.95**	-15.49	-1.99	-5.55**	-12.57	-1.03
$\ln e_{ot}$	-0.64	-2.41	0.89	-0.72	-2.35	0.96

Note: *, ** and *** indicate significance at the 10%, 5% and 1% level respectively.

about future earnings are formed and updated while the significant coefficient on migrant stock indicates network effects. Moreover, it is interesting to note that both levels and first-differences of the explanatory variables turn out significant. This suggests that even though migration is essentially a forward-looking decision, it also strongly fluctuates with short-run cyclical conditions rather than being a steady flow.

With respect to the dynamic specification of the model and the estimation procedure, two points are worth elaborating on. First, note that misspecifying the model especially by omitting lagged migration has a strong impact, most notably on the coefficients of the migrant stock which (looking at the FE estimates) increase from 0.44 (0.24) to 0.73 (0.82) for migration from advanced (developing) countries. Misspecifying the model by omitting the migrant stock results in a less pronounced increase in the coefficient on the lagged migrant flow. Second, correcting for the dynamic panel bias is very important for the coefficient on the lagged migrant flow, which rises from 0.44 (0.61) to 0.61 (0.75) for migration from advanced (developing) countries. Also the coefficients on the other determinants are affected by misspecifying dynamics and/or ignoring the dynamic panel data bias. Especially employment rates in the host country are only then found to be significantly positive for migration from both advanced and developing countries.

All these findings indicate that dynamics play a prominent role in the migration model and should definitely not be ignored, both when specifying the model and selecting the estimation method. Below, we discuss the estimation results for the determinants income and employment separately, focusing on the BCFE estimator.

3.2.2 Income

First, consistent with the findings in the empirical literature (see also Karemera et al., 2000; Mayda, 2010), per capita income in the destination country turns out to be one of the key incentives for migration to OECD countries. For both changes and levels, the coefficient is positive and highly significant across sources of migration. This finding is also robust over the different specifications and estimation methods. Looking at the coefficients on the first-differences, a 1% rise in per capita income in the destination country results in a 1.51% (2.45%) immediate temporary rise in the migrant flow from advanced (developing) countries. The coefficients on the one year lagged per capita income show that this 1% increase attracts an additional 0.59% (1.58%) migrants from advanced (developing) source countries in the next year. In the long run (see Table 2), this amounts to a 1.64% (7.00%) increase in the migrant flow when only taking into account dynamics through the lagged migrant flow (semi long-run effects) and even to a 5.32% (14.82%) increase when also taking into account the link between flows and stocks (full long-run effects). This suggest that taking into account network effects when calculating long-run effects is very important. However, as noted above the exact numbers for the full long-run effects should be considered with care due to the somewhat loose connection between flows and stocks in our dataset.

Second, evidence for the impact of per capita income in the source country is less evident. Both in the short and in the long run, the estimates indicate a statistically significant negative impact on migration for lagged per capita income in advanced origins, but an insignificant impact for per capita income in developing origins (see also Mayda, 2010). First-differenced per capita income at home does not influence the size of migrant flows.

Third, the impact of public services in the destination country is more ambiguous. Immigrants from advanced origin countries prefer destinations with lower levels of public services: the level of public services has a statistically significant elasticity of -0.49 which results in a semi long-run elasticity of -1.19% and a full long-run elasticity of -2.31%. This finding might be explained by the fact that immigrants from advanced countries consider more public services to go together with more social expenditures which can only be financed by higher taxes. In the short run, the level of public services does not appear to have an impact on migration from developing countries, but the immediate response to an increase in public services, as captured by its first-difference, is found significantly positive. In the long run, however, the level of public services does appear significant with the expected positive sign. Immigrants from developing countries may look upon public services as a safety net and move to countries where public services become more generous, in correspondence with the welfare state hypothesis (see also Borjas, 1999).

3.2.3 Employment

Migration from advanced countries seems independent of the actual level of employment rates at home and abroad, and responds only in the short term to changes in the employment opportunities in the host country. In fact, for immigrants from advanced countries the coefficient of changes in the host country's employment rate is the largest of all coefficients. Furthermore, also immigrants from developing countries respond positively to higher employment growth in the destination, though with a smaller and

less significant coefficient. On average, a 1% higher growth in the host country's employment rate results in a temporary increase in the bilateral migrant flow from advanced (developing) countries by 2.18% (1.89%). Against expectations, however, our estimates suggest that migrants from developing countries generally move to countries where employment opportunities are lower. The same result is obtained in the long run, but the coefficient decreases when the link between stocks and flows is accounted for.

4 Conclusions

In this paper we analyze the determinants of international migration to 19 OECD countries from both advanced and developing origin countries between 1998 and 2007 using the OECD's International Migration Database. The contribution of this paper is twofold. First, we estimate a dynamic model of migration based on Hatton's (1995) model using a three-way panel data model. This framework allows to control for observed and unobserved time invariant bilateral effects like geographical, historical, political and cultural influences as well as for time effects like cyclical influences, policy changes, decreases in transportation and communication costs, ..., which are common for all country pairs and reduce the risk of biased results. Including both lagged migration and migrant stocks allows us to separately identify network effects and dynamics stemming from partial adjustment. Second, we estimate this dynamic panel data model using an extended version of the iterative bootstrap algorithm suggested by Everaert and Pozzi (2007). This estimator allows us to correct for the dynamic panel data bias of the FE estimator, which in our model is induced by the lagged migrant flow as well as by the migrant stock, and explicitly takes into account the dynamic relationship between immigrant flows and stocks.

Our results strongly confirm the hypotheses of the human capital theory as well as the network theory of migration, though with a few exceptions. We find that recent immigration to the OECD is primarily driven by better income opportunities in the member states. The influence of income at home and employment rates both at home and abroad is much less pronounced. More specifically, our estimates suggest that immigrants from developing countries are primarily driven by per capita GDP in the host country, whereas variations in migration from advanced countries are determined largely by short-run fluctuations in employment rates abroad. Moreover, as expected, higher native wages in advanced countries seem to discourage immigration, but we find no evidence for an impact of home wages in developing countries.

Furthermore, migrants from advanced countries are unlikely to move to countries with high public services due to the link between social expenditures and tax rates. This is not the case for migrants from developing countries, who consider public expenditures a safety net and prefer countries with rising social expenditures, providing some indication for the welfare magnet hypothesis.

Finally, we find evidence of strong dynamic effects. Both the lagged migration flow and the migrant stock have a strong positive and significant impact on current migration, the former indicating dynamic effects stemming from the process by which expectations about future earnings are formed and updated while the latter indicates network effects. Further evidence that dynamics play a prominent role in the migration model arises from the observation that misspecifying the model by omitting the lagged

migration flow or the migrant stock and/or not correcting for the dynamic panel bias has a strong impact on the estimation results. Therefore, care should be taken when specifying the dynamic structure of the model and selecting the estimation method.

Acknowledgements

We are grateful to Bruno Merlevede for a valuable discussion of a previous version of the paper, to an anonymous referee and the associated editor for constructive comments and suggestions and to participants to the "Seventh GEP Postgraduate Conference" (Nottingham, April 2008) and the "European Trade Study Group Conference" (Warsaw, September 2008) as well as internal seminars at Ghent University, Belgium. Responsibility for any remaining errors lies with the authors. We acknowledge financial support from the Interuniversity Attraction Poles Program - Belgian Science Policy, contract no. P5/21.

References

- Arellano, M. (2003). *Panel data econometrics*. Oxford University Press.
- Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Rev Econ Stud*, 58:277–297.
- Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *J Econom*, 68:29–51.
- Baltagi, B. (2008). *Econometric analysis of panel data*. Wiley, 4 edition.
- Bauer, T. and Zimmermann, K. (1999). Assessment of possible migration pressure and its labour market impact following EU enlargement: enlargement to Central and Eastern Europe. *IZA Research Report*, 3.
- Bertocchi, G. and Strozzi, C. (2008). International migration and the role of institutions. *Public Choice*, 137:81–102.
- Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *J Econom*, 87:115–143.
- Boeri, T., Hanson, G., McCormick, B., and Debenedetti, F. (2002). *Immigration policy and the welfare system: a report for the Fondazione Rodolfo Debenedetti*. Oxford University Press, USA.
- Borjas, G. (1987). Self-selection and the earnings of immigrants. *Am Econ Rev*, 77(4):531–553.
- Borjas, G. (1989). Economic theory and international migration. *Int Mig Rev*, 23(3):457–485.
- Borjas, G. J. (1999). Immigration and welfare magnets. *J Labor Econ*, 17(4):607–37.
- Brücker, H. and Siliverstovs, B. (2006). On the estimation and forecasting of international migration: how relevant is heterogeneity across countries? *Empir Econ*, 31(3):735–754.

- Bun, M. (2003). Bias correction in the dynamic panel data model with a nonscalar disturbance covariance matrix. *Econom Rev*, 22(1):29–58.
- Bun, M. and Carree, M. (2005). Bias-corrected estimation in dynamic panel data models. *J Bus Econ Stat*, 23(2):200–210.
- Bun, M. and Kiviet, J. (2006). The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models. *J Econom*, 132(2):409–444.
- Bun, M. J. G. and Windmeijer, F. (2010). The weak instrument problem of the system GMM estimator in dynamic panel data models. *Econom J*, 13(1):95–126.
- Clark, X., Hatton, T. J., and Williamson, J. G. (2002). Where do U.S. immigrants come from, and why? *NBER Working Paper Series*, w8998.
- Dunlevy, J. A. and Gemery, H. A. (1977). The role of migrant stock and lagged migration in the settlement patterns of nineteenth century immigrants. *Rev Econ Stat*, 59(2):137–44.
- Everaert, G. and Pozzi, L. (2007). Bootstrap-based bias correction for dynamic panels. *J Econ Dyn Control*, 31(4):1160–1184.
- Fertig, M. (2001). The economic impact of EU-enlargement: assessing the migration potential. *Empir Econ*, 26(4):707–720.
- Gallardo-Sejas, H., Gil-Pareja, S., Llorca-Vivero, R., and Martínez-Serrano, J. (2006). Determinants of European immigration: a cross-country analysis. *Appl Econ Lett*, 13(12):769–773.
- Gould, J. (1979). European intercontinental emigration: patterns and causes. *J Eur Econ Hist*, 8:593–679.
- Hatton, T. (1995). A model of UK migration, 1870-1913. *Rev Econ Stat*, 77(3):407–415.
- Hooghe, M., Trappers, A., Meuleman, B., and Reeskens, T. (2008). Migration to European countries: a structural explanation of patterns, 1980-2004. *Int Mig Rev*, 42(2):476–504.
- Karemera, D., Oguledo, V. I., and Davis, B. (2000). A gravity model analysis of international migration to North America. *Appl Econ*, 32(13):1745–1755.
- Kiviet, J. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *J Econom*, 68:53–78.
- Laber, G. (1972). Lagged response in the decision to migrate: a comment. *J Reg Science*, 12(2):307–310.
- Lee, E. (1966). Theory of migration. *Demogr*, 3(1):47–57.
- Lewer, J. and Van den Berg, H. (2008). A gravity model of immigration. *Econ Lett*, 99(1):164–167.
- Mayda, A. (2010). International migration: a panel data analysis of the determinants of bilateral flows. *J Popul Econ*, 23:1249–1274.

- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6):1417–1426.
- Ortega, F. and Peri, G. (2009). The causes and effects of international migrations: evidence from OECD countries 1980-2005. *NBER Working Paper Series*, w14833.
- Pedersen, P. J., Pytlikova, M., and Smith, N. (2008). Selection and network effects - migration flows into OECD countries 1990-2000. *Eur Econ Rev*, 52(7):1160–1186.
- Roodman, D. (2009). A note on the theme of too many instruments. *Oxford B Econ Stat*, 71(1):135–158.
- Todaro, M. (1969). Model of labor migration and urban unemployment in less developed countries. *Am Econ Rev*, 59(1):138–148.
- Vogler, M. and Rotte, R. (2000). The effects of development on migration: theoretical issues and new empirical evidence. *J Popul Econ*, 13(3):485–508.
- Warin, T. and Svaton, P. (2008). European migration: welfare migration or economic migration? *Glob Econ J*, 8(3):1–30.
- Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step gmm estimators. *J Econom*, 126:25–51.
- Ziliak, J. (1997). Efficient estimation with panel data when instruments are predetermined: an empirical comparison of moment-condition estimators. *J Bus Econ stat*, 15(4):419–431.

Appendix A Tables

Table A-1: Total yearly immigrant flows in our sample of destination countries (thousands)

Destination	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	Total
Australia	94.2	101.0	111.3	131.2	121.2	125.9	150.0	167.3	179.8	191.9	1373.8
Austria	59.2	72.4	66.0	74.8	92.6	97.2	108.9	101.5	85.4	92.0	850.0
Belgium	50.7	57.8	57.3	66.0	70.2	68.8	72.4	77.4	83.4	93.4	697.4
Czech Republic	7.9	6.8	4.2	11.3	43.6	57.4	50.8	58.6	66.1	102.5	409.2
Denmark	21.3	20.3	22.9	25.2	22.0	18.7	18.8	20.1	24.0	26.2	219.5
Finland	8.3	7.9	9.1	11.0	10.0	9.4	11.5	12.7	13.9	17.5	111.3
Germany	605.5	673.9	648.8	685.3	658.3	601.8	602.2	579.3	558.5	574.8	6188.4
Hungary	16.1	20.2	20.2	20.3	18.0	19.4	22.2	25.6	19.4	22.6	204.0
Italy	111.0	268.0	271.5	232.8	388.1	353.7	319.3	206.8	181.5	252.4	2585.1
Japan	265.5	281.9	345.8	351.2	343.8	373.9	372.0	372.3	325.6	336.6	3368.6
Korea	211.2	198.3	185.4	172.5	170.9	178.3	188.8	266.3	314.7	317.6	2204.0
Luxembourg	10.6	11.8	10.8	11.1	11.0	12.6	12.2	13.8	13.7	15.8	123.4
Netherlands	81.7	78.4	91.4	94.5	86.6	73.6	65.1	63.4	67.7	80.3	782.7
Norway	26.7	32.2	27.8	25.4	30.8	26.8	27.9	31.4	37.4	53.5	319.9
Portugal	6.5	10.5	15.9	151.4	72.0	31.8	34.1	28.1	22.5	32.6	405.4
Spain	57.2	99.1	330.9	394.0	443.1	429.5	645.8	682.7	803.0	920.5	4805.8
Sweden	35.7	34.6	42.6	44.1	47.6	48.0	47.6	51.3	80.4	99.5	531.4
Switzerland	74.9	85.8	87.4	101.4	101.9	94.0	96.3	94.4	102.7	139.7	978.5
United States	653.2	644.8	841.0	1058.9	1059.4	703.5	957.9	1122.4	1266.3	1052.4	9359.8
Total inflow	2397.4	2705.7	3190.3	3662.4	3791.1	3324.3	3803.8	3975.4	4246.0	4421.8	35518.2

Table A-2: Yearly immigrant flows from our sample of advanced origin countries (thousands)

Destination	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	Total
Australia	41.9	43.4	47.3	54.6	46.5	47.5	58.8	65.0	72.2	76.9	554.1
Austria	8.1	9.2	9.3	12.5	10.3	12.4	15.2	17.2	18.4	20.3	132.9
Belgium	30.3	31.5	32.9	33.0	33.2	33.4	35.8	38.1	40.7	41.6	350.5
Czech Republic	2.5	2.0	1.2	2.7	14.5	25.4	17.0	12.9	9.4	17.5	105.1
Denmark	7.8	7.3	7.4	7.4	7.2	7.3	7.8	8.2	9.8	9.8	80.0
Finland	1.4	1.3	1.3	1.4	1.3	1.4	1.5	1.6	1.7	1.9	14.8
Germany	182.7	188.2	187.6	179.5	168.4	152.2	148.0	143.3	148.3	150.7	1648.9
Hungary	2.5	2.9	3.3	2.9	2.4	2.7	1.6	9.0	4.1	3.0	34.4
Italy	6.1	7.8	9.0	9.2	13.2	11.8	10.3	8.0	6.1	5.1	86.7
Japan	51.6	54.8	55.3	52.0	51.0	50.0	50.4	51.1	53.5	56.7	526.4
Korea	17.3	19.6	21.9	24.2	27.5	24.4	25.4	27.4	27.2	28.8	243.7
Luxembourg	7.9	8.2	8.5	8.7	8.4	9.3	9.0	9.9	10.3	11.1	91.3
Netherlands	27.0	27.7	29.9	30.2	28.3	25.7	24.7	24.8	29.0	31.1	278.4
Norway	15.9	12.7	11.4	11.5	11.4	9.9	10.0	10.6	12.9	16.6	122.9
Portugal	2.9	4.3	4.4	4.9	4.3	3.9	4.0	3.5	2.4	9.9	44.5
Spain	24.2	35.3	45.6	55.7	72.4	76.6	99.0	109.8	124.3	133.0	775.9
Sweden	9.9	11.1	13.7	14.5	15.4	15.2	14.6	14.9	17.7	16.5	143.5
Switzerland	39.2	44.2	47.2	50.3	54.1	53.9	59.3	60.1	66.7	98.0	573.0
United States	66.5	58.9	89.0	119.6	107.2	66.7	95.8	124.3	113.3	96.8	938.1
Total inflow	479.2	511.5	537.2	555.2	569.8	563.0	592.4	615.4	654.7	728.5	5807.0
% of total inflow	19.99	18.90	16.84	15.16	15.03	16.94	15.57	15.48	15.42	16.48	16.35

Note: % of total inflow denotes the share of yearly immigrant flows from advanced origin countries covered in total yearly immigrant flows in the destination countries.

Table A-3: Yearly immigrant flows from our sample of developing origin countries (thousands)

Destination	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	Total
Australia	44.1	48.2	54.3	64.3	63.9	69.4	81.1	87.7	92.8	104.2	710.0
Austria	20.3	24.3	23.5	28.8	28.0	29.7	32.7	31.2	25.1	29.7	273.3
Belgium	12.2	13.3	15.5	21.4	24.4	23.2	24.4	26.2	29.7	37.5	227.8
Czech Republic	3.9	3.5	2.3	7.0	22.9	25.3	27.2	37.0	46.5	67.4	243
Denmark	9.4	9.0	10.7	13.0	10.1	7.6	6.6	6.6	7.6	7.7	88.3
Finland	4.8	4.1	4.6	5.7	5.3	4.8	5.7	6.5	7.1	8.9	57.5
Germany	268.5	298.5	329.9	359.0	357.5	329.7	342.2	337.1	320.3	341.3	3284
Hungary	10.5	12.9	13.4	14.4	13.4	13.9	17.8	12.5	12.0	13.1	133.9
Italy	80.8	176.2	184.3	163.8	252.1	235.2	224.0	139.2	124.0	217.8	1797.4
Japan	144.3	160.9	215.3	223.1	219.4	243.3	243.0	232.8	196.4	202.5	2081
Korea	9.3	8.6	7.9	7.2	10.0	9.3	5.2	10.3	6.8	5.2	79.8
Netherlands	29.2	25.4	27.1	30.1	31.4	32.2	29.0	28.7	29.9	37.9	300.9
Norway	9.8	11.7	13.9	11.5	16.5	14.2	13.7	16.6	20.5	31.9	160.3
Portugal	2.3	4.3	8.4	53.2	29.6	14.6	20.8	16.1	11.9	12.4	173.6
Spain	31.7	60.7	276.5	328.2	364.2	347.9	366.3	451.4	530.9	642.9	3400.7
Sweden	16.7	15.1	17.5	17.9	20.7	20.3	19.3	19.9	37.8	42.1	227.3
Switzerland	3.9	4.7	4.5	5.3	5.3	4.6	4.4	4.1	4.7	4.7	46.2
United States	527.7	529.8	676.1	833.7	838.9	571.2	773.9	875.1	1014.0	850.4	7490.8
Total inflow	1229.4	1411.2	1885.7	2187.6	2313.6	1996.4	2237.3	2339.0	2518.0	2657.6	20775.8
% of total inflow	51.28	52.16	59.11	59.73	61.03	60.05	58.82	58.84	59.30	60.10	58.49

Note: % of total inflow denotes the share of yearly immigrant flows from developing origin countries covered in total yearly immigrant flows in the destination countries.

Table A-4: Descriptive statistics

Variable	Advanced origin countries				Developing origin countries			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
$\ln M_{dot}$	6.82	1.49	4.61	10.70	7.06	1.77	4.61	12.30
$\ln MST_{dot}$	9.79	1.71	5.80	14.14	1.00	1.83	5.58	16.46
$\ln M_{dot-1}$	6.78	1.48	4.61	10.70	6.97	1.76	4.61	12.30
$\ln w_{dt-1}$	10.27	0.31	9.22	11.29	10.25	0.25	9.22	10.84
$\ln w_{ot-1}$	10.17	0.28	9.19	11.29	8.41	0.84	5.66	10.89
$\ln ps_{dt-1}$	2.08	0.18	0.75	2.41	2.07	0.15	0.75	2.33
$\ln e_{dt-1}$	4.04	0.12	3.74	4.20	4.04	0.12	3.74	4.20
$\ln e_{ot-1}$	4.02	0.11	3.74	4.31	4.00	0.19	3.41	4.45
$\ln \Delta w_{dt}$	0.05	0.02	-0.01	0.16	0.05	0.02	-0.01	0.16
$\ln \Delta w_{ot}$	0.05	0.02	-0.07	0.16	0.06	0.07	-0.50	0.59
$\ln \Delta ps_{dt}$	0.01	0.04	-0.11	0.25	0.01	0.03	-0.11	0.25
$\ln \Delta e_{dt}$	0.00	0.01	-0.02	0.05	0.01	0.01	-0.02	0.05
$\ln \Delta e_{ot}$	0.00	0.01	-0.04	0.05	0.00	0.03	-0.15	0.22

Note: *Advanced*: 2223 observations, 247 cross sections. *Developing*: 3492 observations, 388 cross sections.

Table A-5: GMM estimation results

Dependent variable: $\ln M_{dot}$

Sample period: 1998-2007

	Advanced origins				Developing origins			
	GMMd1	GMMd2	GMMs1	GMMs2	GMMd1	GMMd2	GMMs1	GMMs2
$\ln MST_{dot}$	1.57*** (4.56)	1.08** (2.44)	0.44*** (8.35)	0.34*** (3.12)	-0.01 (-0.09)	-0.10 (-0.46)	0.92*** (43.26)	0.94*** (22.37)
$\ln M_{dot-1}$	0.98*** (7.46)	1.01*** (5.36)	0.51*** (10.76)	0.60*** (5.74)	8.39*** (4.75)	8.57 (1.49)	-0.07 (-0.51)	0.08 (0.33)
$\ln w_{dt-1}$	-2.52** (-2.44)	-1.52 (-1.01)	-0.36** (-2.22)	-0.23 (-1.14)	0.42 (1.05)	0.31 (0.28)	0.12** (2.45)	0.09 (1.08)
$\ln w_{ot-1}$	0.99 (0.84)	0.59 (0.40)	0.00 (-0.01)	-0.02 (-0.08)	4.68*** (2.91)	4.53 (0.76)	-0.17 (-1.36)	-0.54** (-2.49)
$\ln ps_{dt-1}$	-1.09 (-0.57)	0.79 (0.32)	-0.03 (-0.26)	-0.48** (-2.15)	9.55*** (4.32)	10.57** (2.21)	-0.27 (-1.26)	-0.16 (-0.41)
$\ln e_{dt-1}$	-1.33 (-1.21)	-2.12 (-1.61)	0.17 (0.55)	0.51 (1.08)	15.45*** (4.33)	16.67** (2.00)	0.25 (1.51)	0.06 (0.22)
$\ln e_{ot-1}$	2.80* (1.84)	1.69 (0.85)	0.50** (2.32)	0.21 (0.43)	-1.36 (-1.27)	-1.68 (-0.77)	1.16** (2.48)	1.03 (1.40)
$\ln \Delta w_{dt}$	2.39*** (2.83)	2.18** (2.27)	0.88* (1.79)	0.33 (0.67)	0.34* (1.67)	0.31 (0.92)	0.08 (0.87)	0.04 (0.27)
$\ln \Delta w_{ot}$	-1.28 (-1.31)	-0.20 (-0.17)	-0.73* (-1.71)	-0.22 (-0.52)	-1.27* (-1.77)	-1.19 (-0.93)	-0.13 (-0.39)	-0.34 (-0.65)
$\ln \Delta ps_{dt}$	0.55 (0.69)	0.03 (0.03)	0.71** (2.05)	0.49 (1.43)	-0.63 (-0.27)	-0.82 (-0.19)	5.35*** (5.29)	4.49*** (2.85)
$\ln \Delta e_{dt}$	3.52** (2.48)	2.96** (2.20)	1.33 (1.18)	0.90 (1.33)	1.12** (2.15)	0.95 (0.96)	0.07 (0.23)	-0.22 (-0.41)
$\ln \Delta e_{ot}$	-2.98 (-1.60)	-1.83 (-0.76)	0.35 (0.36)	-0.01 (-0.01)	-0.16** (-2.37)	-0.15 (-0.57)	0.07*** (3.08)	0.06* (1.70)
<i>Sargan and Hansen tests of overidentifying restrictions</i>								
χ^2	2.28	2.05	13.65	43.93	3.68	3.77	33.36	90.73
p-value	0.52	0.56	0.55	0.00	0.30	0.29	0.00	0.00

Notes: Each regression includes time dummies (not reported). t -statistics - between brackets - are robust to cross-sectional heteroskedasticity. For the two-step GMM estimators they are calculated from Windmeijer (2005) standard errors. *, ** and *** indicate significance at the 10%, 5% and 1% level respectively. *Advanced*: 2223 observations and 247 cross sections. *Developing*: 3492 observations and 388 cross sections.

The following estimators are reported: (i) GMMd1 and GMMd2, the stacked one-step and two-step versions of the first-differenced GMM estimator proposed by Arellano and Bond (1991) and (ii) GMMs1 and GMMs2, the stacked one-step and two-step versions of the system GMM estimator proposed by Arellano and Bover (1995) and Blundell and Bond (1998). One-step GMM estimators report the Sargan test of overidentifying restrictions; two-step variants report the robust Hansen test. The instrument sets for GMMd and GMMs are exactly the same as the ones used in the Monte Carlo simulation in section 2.5.